Aliasing artefact index for image interpolation quality assessment

Rukundo, Olivier; Schmidt, Samuel

Published in:
Optoelectronic Imaging and Multimedia Technology V

DOI (link to publication from Publisher):
10.1117/12.2503872

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Aliasing artefact index for image interpolation quality assessment

Olivier Rukundo, Samuel E. Schmidt

Olivier Rukundo, Samuel E. Schmidt, "Aliasing artefact index for image interpolation quality assessment," Proc. SPIE 10817, Optoelectronic Imaging and Multimedia Technology V, 108171E (7 November 2018); doi: 10.1117/12.2503872

SPIE.

Event: SPIE/COS Photonics Asia, 2018, Beijing, China

Aliasing artefact index for image interpolation quality assessment

Olivier Rukundo[1a], Samuel E. Schmidt[a]
[a]Department of Health Science and Technology, Fredrik Bajers Vej 7, 9220 Aalborg, Denmark

ABSTRACT

A preliminary study of a non-reference aliasing artefact index (AAI) metric is presented in this paper. We focus on the effects of combining a full-reference metric and interpolation algorithm. The nearest neighbor algorithm (NN) is used as the gold standard against which test-algorithms are judged in terms of aliased structures. The structural similarity index (SSIM) metric is used to evaluate a test image (i.e. a test-algorithm's image) and a reference image (i.e. the NN's image). Preliminary experiments demonstrated promising effects of the AAI metric against state-of-the-art non-reference metrics mentioned. A new study may further develop the studied metric for potential applications in image quality adaptation and/or monitoring in medical imaging.

Keywords: aliasing artefact index, nearest neighbor algorithm, image interpolation, objective image quality metric, structural similarity index, gold standard, test-algorithm

1. INTRODUCTION

In image interpolation, the need for quality assessment has become greater due to easy degradation of image interpolation quality and wide use of interpolation algorithms in many engineering areas/fields [7], [8], [9], [10]. Current methods for image quality assessment include *subjective image quality assessment* (which requires human judges or observers) and *objective image quality assessment* (which does not require human judges or observers) [23]. Referring to ITU-R BT.500 recommendations, subjective experiments require *much time* (much higher, if any desired scaling or interpolation ratios is considered) and *careful planning* to finally obtain the individual mean opinion score (MOS), which can be used as a ground truth for the development of objective image quality metrics [1], [2], [23]. Objective image quality metrics can be understood as a speedier scientific solution helping to guess automatically and accurately the MOS [2], [23]. In some works, objective quality methods are classified into psychophysical and engineering approaches and further into full-reference, reduced-reference and non-reference [24]. It is important to note that objective quality assessment methods were or are developed for different purposes, such as to quantify distortions, produce benchmarks, monitor quality, optimize a process, or indicate problem areas [28]. The difference, in purposes, pushed many researchers to put efforts in seeking first to classify such metrics into groups (for further research, etc.) but they were unable to draw a clear boundary between the different groups [28]. The lack of clarity, on such a boundary, triggered a kind of 'classification competition' on quality assessment metrics [31], [32], [33], [34]. For example, in full-reference (FR) metrics, as stated in [28], [29] authors divided such image quality (IQ) metrics into six groups based on the information they use (such as pixel difference-based, correlation-based measures, edge-based measures, spectral distance-based measures, context-based, Human Visual System (HVS)-based). In [30] authors divided IQ metrics into two distinct groups (based on HVS model for low-level perception and prior knowledge about introduced distortion). In [31] authors classified IQ metrics based on three criteria (i.e. full-reference, non-reference, and reduced-reference metrics; general-purpose and application-specific metrics and; bottom-up and top-down metrics). In [28] authors divided the IQ metrics into four groups and giving some of their examples (such as PSNR, MSE, CSF, VSNR, SSIM, etc.). Now, in non-reference (NR) metrics classification, there are, even, structured categories and subcategories based on the types of methodologies used for image quality estimation. Some categories include pixel-based methods, bitstream-based methods and hybrid methods combining the previous two. For example, in [23], authors divided the NR metrics into non-reference pixel based (NR-P), non-reference bitstream based (NR-B), and a hybrid of NR-P and NR-B as well as their subcategories, in each case. It is important to note that the NR-B encompasses the approaches termed as feature-based. And, according to [23], the feature-based methods are either based on a specific artefacts model (related to a visible degradation, the same as in this AAI metric study) or a model for the quantification of degradation's effects in pristine images or videos [25]. NR methods are discussed again in [26], under the categorization of features and artifacts detection [23]. Again, NR methods have been reviewed and classified following approaches such as an approach based

[1]orukundo@hst.aau.dk; phone 45 4294 8142; hst.aau.dk

Optoelectronic Imaging and Multimedia Technology V, edited by Qionghai Dai, Tsutomu Shimura, Proc. of SPIE Vol. 10817, 108171E · © 2018 SPIE
CCC code: 0277-786X/18/$18 · doi: 10.1117/12.2503872

on the type of distortion (that can be estimated to formulate a quality value) and quantification of artifacts produced (using a specific compression standard) [27], [23]. Despite the current several categories of FR and NR methods [19], [20], [21], [22], further studies are still needed, particularly on image interpolation quality assessment, still in its infancy [15]. It is important to note that in objective image interpolation quality assessment, FR methods requires inferring pristine images [14], [15], [17] while the NR methods require the use of additional complex functions to assess aliasing artefacts or distortions [20], [21]. Also, it is important to note that such aliasing distortions are the most rapidly visible and poorest-quality-prone particularly at higher scaling or interpolation ratios [3], [4], [5], [6]. This paper is organized as follows: Part two briefly introduces the NN algorithm and SSIM. Part 3 presents and explains the working principle of the studied AAI metric. Part four presents preliminary experiments conducted. And, the conclusion is given in part five.

2. NEAREST NEIGHBOR ALGORITHM AND STRUCTURAL SIMILARITY INDEX

In the current literature, there exist many interpolation algorithms developed for different purposes [4], [5], [13], [16], [18], but here, a great interest has been put on the NN algorithm. Equation 1 is the basic linear scaling equation on which the NN interpolation algorithm is based. In Equation 1, the source or original image coordinates are represented by *scrCoord*, destination or interpolated image coordinates by *destCoord*, source image length by *srcLength* and destination image length by *destLength*.

$$scrCoord = \left(\frac{destCoord}{\left(\frac{destLength}{scrLength} \right)} \right) \tag{1}$$

Figure 1 illustrates the source and destination elements mentioned above.

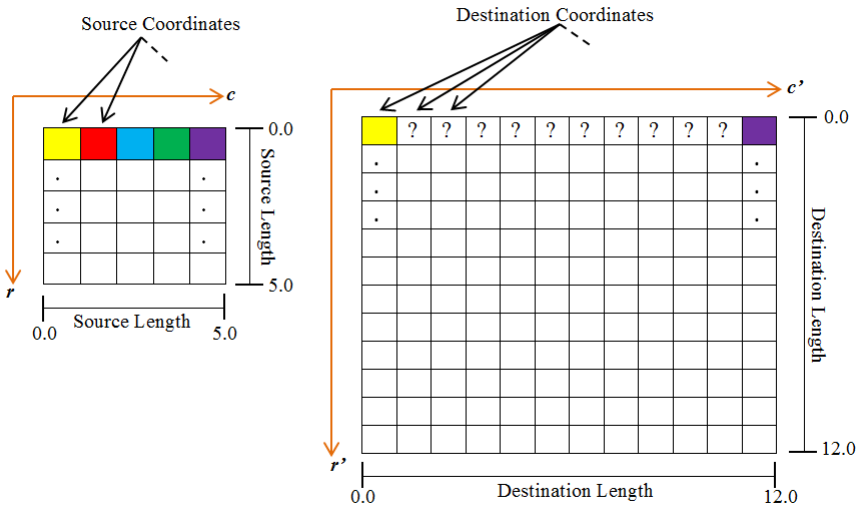

Figure 1: Here, c, c' and r, r' represent the source and destination image coordinates, in the horizontal and vertical directions, respectively.

In the NN algorithm, the rounding function is generally used to keep to the integerness requirement of the digital image formats [11], and to select the nearest pixel, by turning any non-integer, producible by Equation 1, to an integer value, that represents the value for next nearest pixel (to that non-integer value). In general, functions used for rounding purposes, such as floor $\lfloor\ \rfloor$, round $[\ ]$, ceil $\lceil\ \rceil$ functions, are based on Equation 2.

$$\begin{aligned}\lfloor x \rfloor &= x - (x \bmod 1)\\ [x] &= (x + 0.5) - ((x + 0.5) \bmod 1)\\ \lceil x \rceil &= (x + 1) - ((x + 1) \bmod 1)\end{aligned} \tag{2}$$

The structural similarity index (SSIM) is one of image quality assessment (IQA) metrics that brought IQA metrics from pixel-to structure-based stage [22]. In [28], authors categorized the SSIM as a high-level metric. The SSIM assumes that human visual perception is highly adapted for extracting structural information from a scene, and it defines the structural information as those attributes that represent the structure of objects in the scene, and that are independent of the average luminance and contrast [19]. The SSIM index between two nonnegative images, of the same size, is given by Equation 3.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3}$$

where, x and y are two nonnegative image signals, μ_x is the mean intensity, σ_x^2 is the variance of x, σ_x is the covariance of x, C_1 and C_2 are variables stabilizing the division with the weak denominator, for example, when the sum of variances is very close to zero [19].

3. THE STUDIED AAI METRIC

The studied AAI metric involves three main parts, namely, the image scaling down part, the image upscaling part using the NN algorithm and SSIM metric (see Figure 2). In this study, a grayscale image is selected as a test image. The scaling down is achieved using the method proposed in [15], but the use of other methods is possible to investigate. A test-ratio is also used (i.e. the scaling ratio on which will be used during the interpolation operation) and is equivalent to the inverse of the input test ratio(for the scaling down purposes) otherwise, it is originally for NN upscaling purposes. The NN algorithm is used to scale up the image scaled down in the previous part. It is important to note that the choice of NN algorithm is because the NN algorithm does not create new image pixel values, even, at different scaling or interpolation ratios. In other words, the NN algorithm does not change the entropy of the test image during image interpolation operations, which is good for our study. The choice of SSIM is because it uses structural information from a test image. On top of that, such structural information is very important to judge test-algorithms against our aliasing artefact based gold standard (i.e. the NN algorithm). The NN interpolated and test images are used as inputs to the SSIM part. And, the NN image is used as a reference image, while the test image is used as in input image.

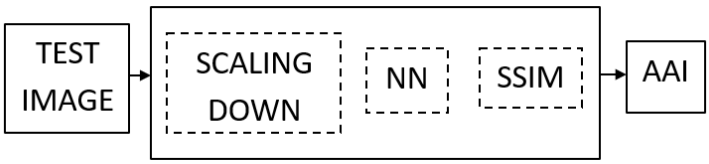

Figure 2: A functional block diagram of the studied AAI metric

The value of AAI obtained ranges between 0 and 1. Traditionally, if after interpolation of the test image, a high SSIM value (i.e. very close or equal to 1) is obtained, it means that the test image is closely related to the pristine or reference image. Here, if this happens, it means that information contained in the test image is like (or almost like) the information contained in the NN interpolated image, which would not be good since the NN algorithm produces the heaviest aliased structures after interpolation. Since, in the traditional context, a higher SSIM values means a higher similarity to the reference (i.e. better image quality), and, since the reference in our case, is the image interpolated using (or obtained thanks to) the NN algorithm, it can be understood that a higher AAI value would simply reflect the undesired or unwanted similarity, thus making the smaller AAI value to reflect the wanted dissimilarity in terms of aliasing or blocking artefacts. Preliminary experiments are presented and discussed in the following part.

4. EXPERIMENTS

The preliminary AAI algorithm has been implemented in MATLAB-R2018a. Objective assessments were conducted using grayscale images downloaded from the USC-SIPI Image database [12]. Test algorithms (or interpolation methods) used for comparisons include the bilinear, bicubic and lanczos. Traditionally, assessments based on subjective and objective assessments are conducted to determine the correlation characteristics between each objective metric and subjective assessment [35]. Here, we first look at the behaviors of the AAI metric against BRISQUE and NIQE metrics, at various scaling ratios, as shown in Figure 3, Figure 4, Figure 5, Figure 6, Figure 7 and Figure 8. Further, we use the correlation between the AAI and each of the two methods, separately, as shown in Table 1.

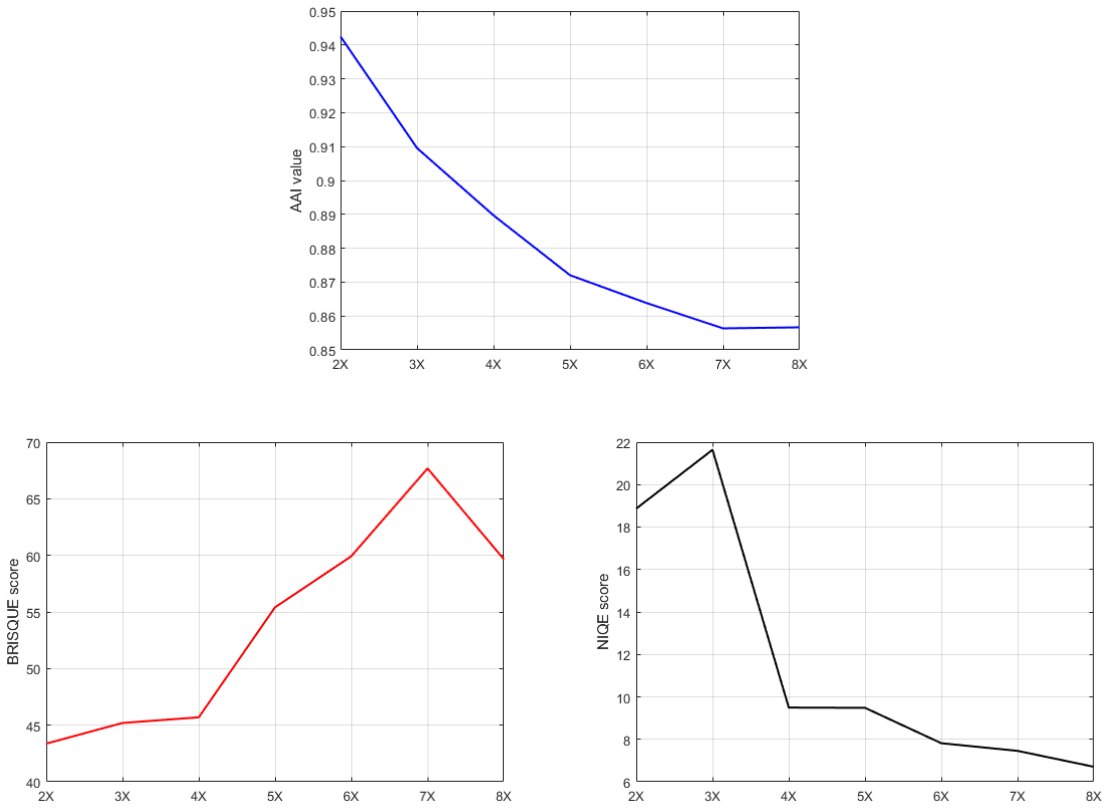

Figure 3: With bilinear test-algorithm and Lenna input image size 64 x 64: AAI (top), BRISQUE (left) and NIQE (right)

Normally, as the interpolation ratio increases, the image interpolation algorithm produces more artefacts, particularly the jaggedness or aliasing artefacts. Aliasing artefacts are the easiest artefact to see, with naked eyes. In our case, as the scaling ratio changes from two times to eight times, (i.e. 2X to 8X), the AAI demonstrates a strong dissimilarity as shown by quasi-linearly decreasing values (see all figures of this part), while the BRISQUE produces scores go up (which means a strong presence of distortions) but not the same way as in the AAI. Also, the NIQE produces values that go down (meaning a strong presence of distortions) but not the same way as in the AAI (and, mostly for the case involving images of the size 64 x 64). In Table 1, the first column, the first group of images of size = 64 x 64, there is a strong negative correlation between the AAI and BRISQUE, and a strong positive correlation between the AAI and NIQE scores, for cases involving Pepper image as well as bilinear, bicubic and lanczos test-algorithms. In the same table's column, but in the second group of images of size = 128 x 128, there is a strong negative correlation between the AAI and BRISQUE and a weaker negative correlation between the AAI and NIQE scores, for cases involving Lenna image as well as bilinear, bicubic and lanczos test-algorithms. Still in the same table's column, but in the third group of images of size = 256 x 256, there is a strong negative correlation between the AAI and BRISQUE scores and a strong correlation between the AAI and NIQE scores, for cases involving Lenna image as well as bilinear, bicubic and lanczos interpolation algorithms.

Table 1: Correlation coefficients

	LENNA (64 x 64)		PEPPERS (64 x 64)	
METHOD	BRISQUE	NIQE	BRISQUE	NIQE
Bilinear	-0.8765	0.8795	-0.8451	0.8255
Bicubic	-0.9722	0.8390	-0.9736	0.7267
Lanczos	-0.9814	0.8379	-0.9686	0.7368
	LENNA (128 x 128)		PEPPERS (128 x 128)	
METHOD	BRISQUE	NIQE	BRISQUE	NIQE
Bilinear	-0.9183	-0.1644	-0.9033	-0.1814
Bicubic	-0.9819	-0.8653	-0.9823	-0.7933
Lanczos	-0.9909	-0.6698	-0.9973	-0.6303
	LENNA (256 x 256)		PEPPERS (256 x 256)	
METHOD	BRISQUE	NIQE	BRISQUE	NIQE
Bilinear	-0.9199	-0.5727	-0.9060	-0.3346
Bicubic	-0.9661	-0.9307	-0.9903	-0.8622
Lanczos	-0.9744	-0.8935	-0.9941	-0.8140

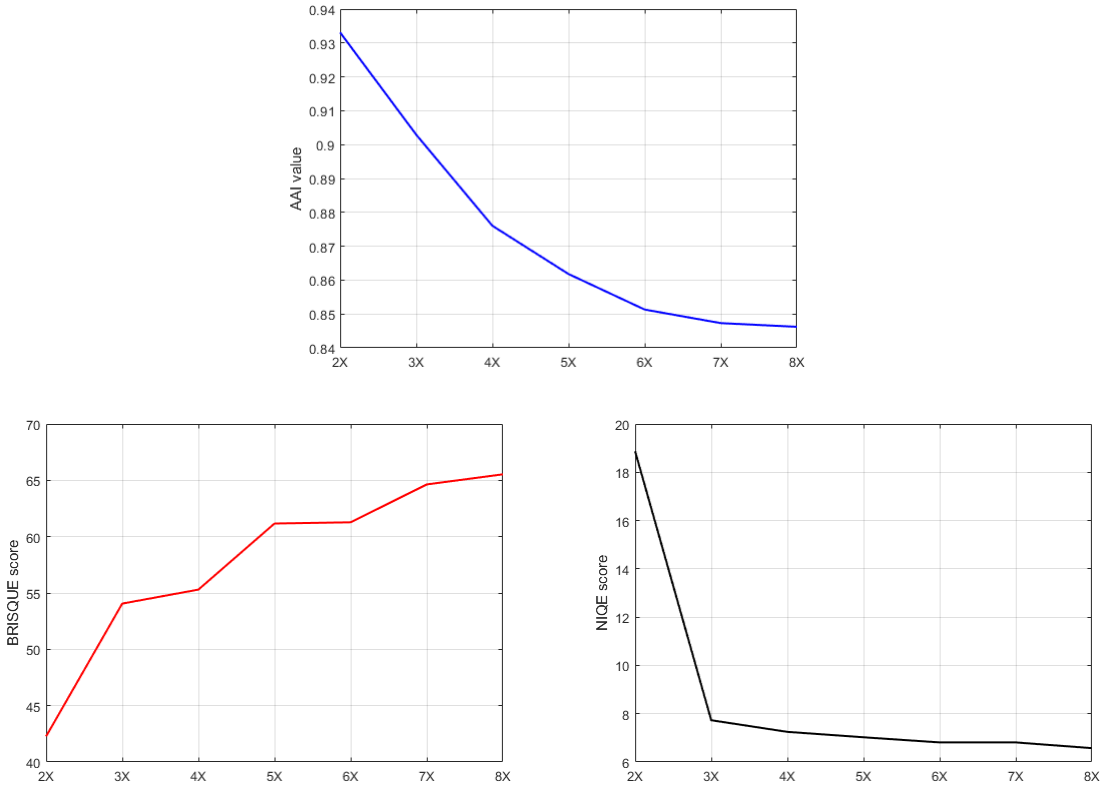

Figure 4: With bicubic test-algorithm and Lenna input image size 64 x 64: AAI (top), BRISQUE (left) and NIQE (right)

Now, in same table's second column, but in the first group of images of size = 64 x 64, there is a strong negative correlation between the AAI and BRISQUE, and a strong positive correlation between the AAI and NIQE scores, for cases involving Pepper image as well as bilinear, bicubic and lanczos test-algorithms. However, in same table's second column, but in the second group of images of size = 128 x 128, there is a strong negative correlation between the AAI and BRISQUE, and a weaker negative correlation between the AAI and NIQE scores, for cases involving Pepper image as well as bilinear, bicubic and lanczos interpolation algorithms. Finally, in same table's second column, but in the third group of images of size = 256 x 256, there is a strong negative correlation between the AAI and BRISQUE, and a weaker negative correlation between the AAI and NIQE scores, for cases involving Pepper image as well as bilinear, bicubic and lanczos test-algorithms.

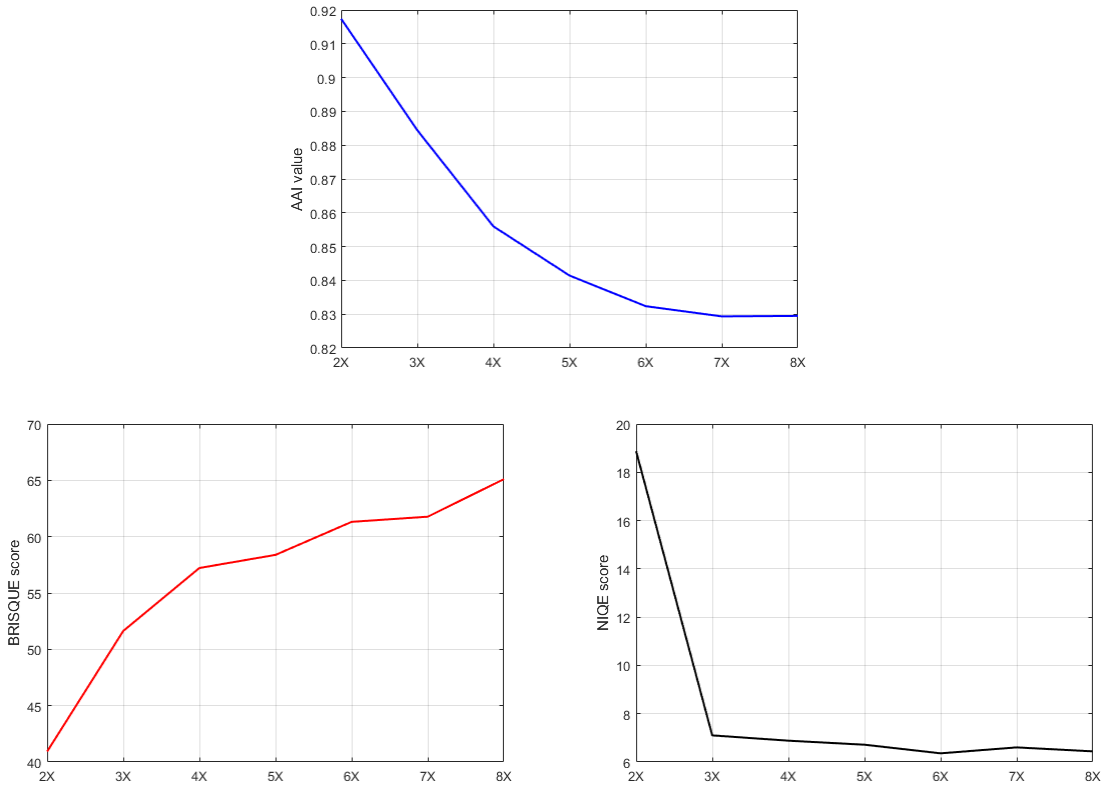

Figure 5: With lanczos test-algorithm and Lenna input image size 64 x 64: AAI (top), BRISQUE (left) and NIQE (right)

5. CONCLUSION

A preliminary assessment method has been studied and presented in this paper. Current methods for image quality assessment include subjective image quality assessment and objective image quality assessment. Recent research efforts focused on the development of FR and NR methods for the image quality assessment purposes. Our study focused on the use of the FR method and interpolation algorithm to achieve a novel NR category method for image interpolation quality assessment. The studied AAI uses the NN algorithm (regarded as the gold standard against which test-algorithms are judged) and the SSIM metric to evaluate the images generated by a test-algorithm and NN algorithm, respectively. The value of AAI ranges from zero to one. The smaller AAI value reflected the wanted dissimilarity, in terms of aliasing or blocking artefacts. Our experiment focused on the behaviors of the AAI metric against the BRISQUE and NIQE metrics, at various scaling ratios and test-algorithms. Further experiments demonstrated that the AAI metric was independent of image size and scaling ratio, demonstrated a strong and constant dissimilarity (with quasi-linearly decreasing values) which was not the case with NIQE as well as BRISQUE. Future research efforts may be dedicated to further development and assessment of the studied metric for potential applications in image quality adaptation and/or monitoring in medical imaging.

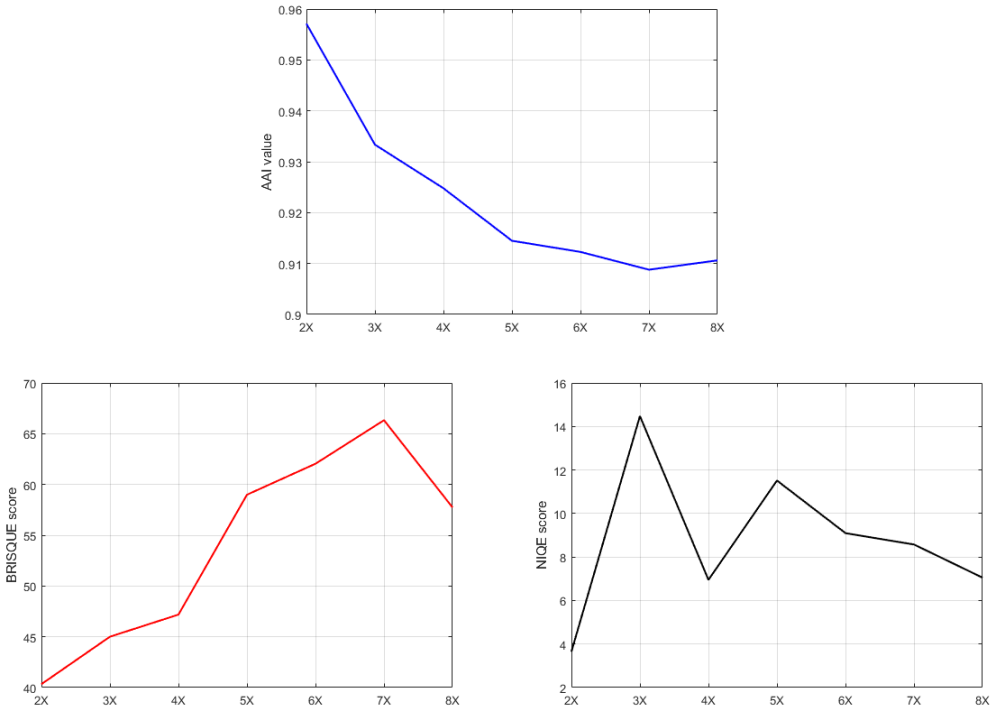

Figure 6: With bilinear test-algorithm and Pepper input image size 256 x 256: AAI (top), BRISQUE (left) and NIQE (right)

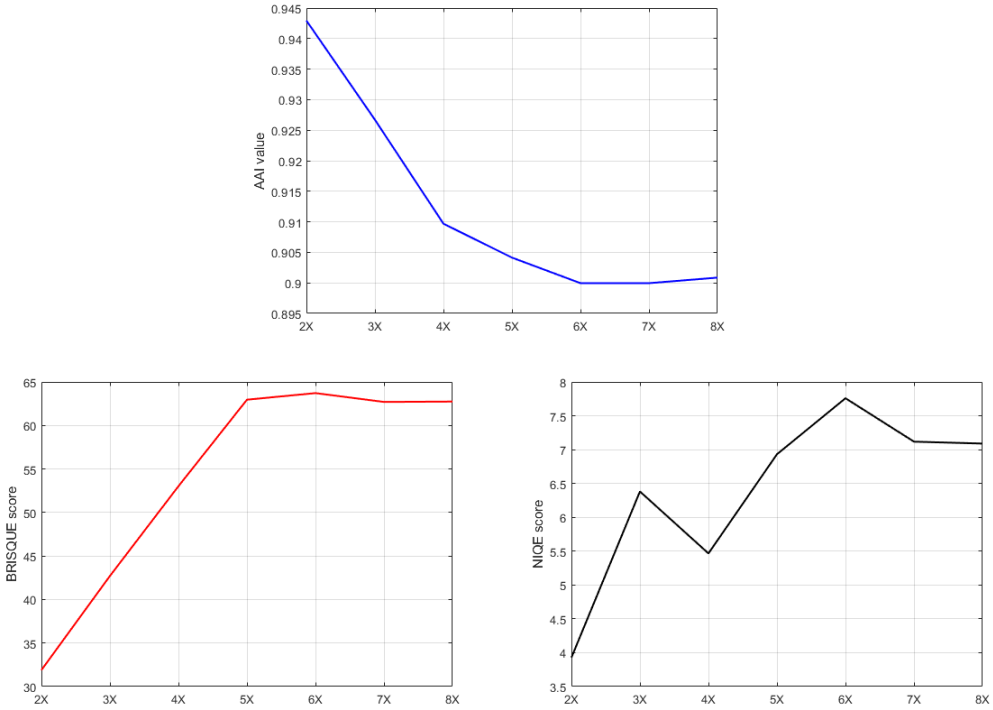

Figure 7: With bicubic test-algorithm and Pepper input image size 256 x 256: AAI (top), BRISQUE (left) and NIQE (right)

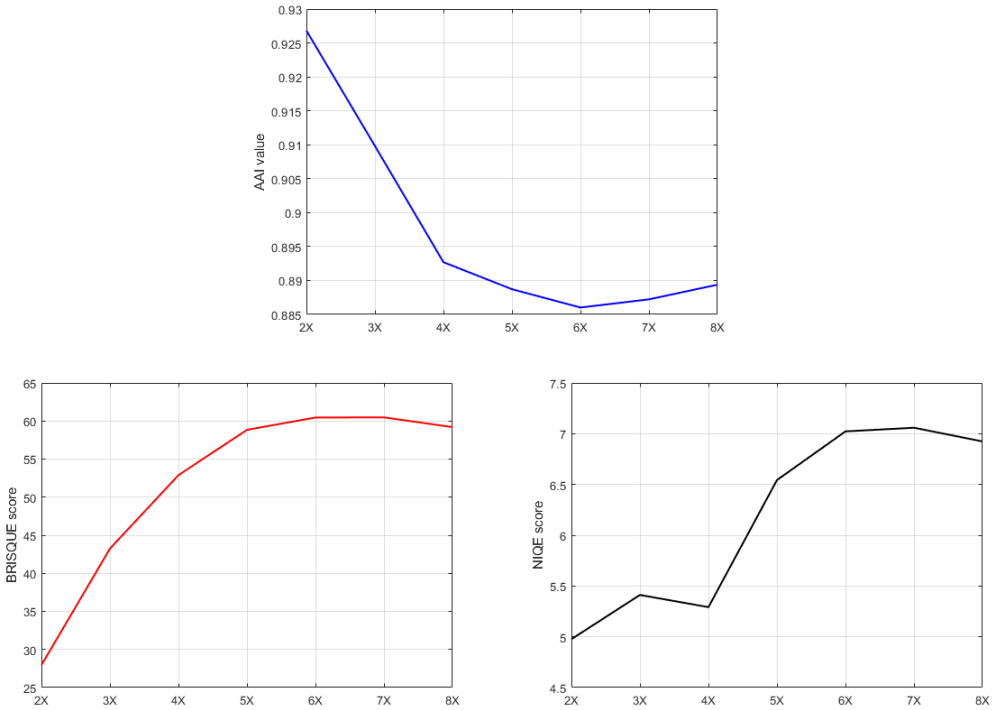

Figure 8: With lanczos test-algorithm and Pepper input image size 256 x 256: AAI (top), BRISQUE (left) and NIQE (right)

REFERENCES

[1] ITU-R Recommendation BT.500-13 (01/2012), "Methodology for the subjective assessment of the quality of the television pictures,"<https://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.500-13-201201-I!!PDF-E.pdf>. Accessed 10 April (2018)

[2] Shahid, M., Rossholm, A., et al, "No-reference image and video quality assessment: a classification and review of recent approaches," EURASIP Journal on Image and Video Processing, 2014:40, (2014)

[3] Rukundo, O., "Thinning based Antialiasing Approach for Visual Saliency of Digital Images," Proceedings of the 10th International Conference on Computer Vision Theory and Applications, pp. 658-665, (2015)

[4] Rukundo, O., Cao, H.Q., "Advances on image interpolation based on ant colony algorithm," SpringerPlus, 5:304, (2016)

[5] Rukundo, O., Maharaj, B.T., "Optimization of image interpolation based on nearest neighbor algorithm," Proceedings of the 9th International Conference on Computer Vision Theory and Applications (VISAPP), pp. 641–647, (2014)

[6] Rukundo, O., Wu, K.N., Cao, H.Q., "Image interpolation based on the pixel value corresponding to the smallest absolute difference. Proceedings of 4th International Workshop on Advanced Computational Intelligence (IWACI), pp. 432–435, (2011)

[7] Sina, M.I., Cretu, A.-M., Payeur, P., "Biological visual attention guided automatic image segmentation with application in satellite imaging," Proc. SPIE 8291, Human Vision and Electronic Imaging XVII, 82911N, (2012)

[8] Rukundo, O., Cao, H.Q., Huang, M.H., "Optimization of bilinear interpolation based on ant colony algorithm," Lecture Notes in Electrical Engineering 137, Springer Berlin Heidelberg, pp. 571–580, (2012)

[9] Herry, C.L., Goubran, R.A., Frize, M., "Improving the Detection and Localization of Anatomical Landmark Points in Infrared Images Using Symmetry and Region Specific Constraints," Proceedings of the 25th IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Victoria, (2008)

[10] Rukundo, O., "Half-unit weighted bilinear algorithm for image contrast enhancement in capsule endoscopy," Proc. SPIE 10615, Ninth International Conference on Graphic and Image Processing,106152Q, (2018)

[11] Rukundo, O., "Effects of improved-floor function on the accuracy of bilinear interpolation algorithm," Computer and Information Science, 8(4), pp. 1–11, (2015)
[12] The USC-SIPI Image Database, USC Viterbi School of Engineering, <http://sipi.usc.edu/database/>, accessed 15/07/2018, (2018)
[13] Li, X., Orchard, M.T., "New edge-directed interpolation," IEEE Trans Image Process, 10(10), pp.1521–1527, (2001)
[14] Rukundo, O., Cao, H.Q., "Nearest neighbor value interpolation," International Journal of Advanced Computer Science and Applications 3(4), pp. 25–30, (2012)
[15] Rukundo, O., Schmidt, S., " Effects of rescaling bilinear interpolant on image interpolation quality," Proc. SPIE 10817, Optoelectronic Imaging and Multimedia Technology V, (2018)
[16] Rukundo, O., Pedersen, M., Hovde, Ø., "Advanced Image Enhancement Method for Distant Vessels and Structures in Capsule Endoscopy," Computational and Mathematical Methods in Medicine, 2017, (2017)
[17] Rukundo, O., Schmidt, S., "Extrapolation for image interpolation," Proc. SPIE 10817, Optoelectronic Imaging and Multimedia Technology V, (2018)
[18] Rukundo, O., "Effects of empty bins on image upscaling in capsule endoscopy," Proc. SPIE 10420, Ninth International Conference on Digital Image Processing, 104202P, (2017)
[19] Wang, Z., Bovik, A.C., Sheikh, H.R., "Image Quality Assessment: From Error Visibility to Structural Similarity", IEEE Transactions on Image Processing, 13(4), pp. 600-612, (2004)
[20] Mittal, A., Moorthy, A. K., Bovik, A. C., "No-reference image quality assessment in the spatial domain," IEEE Transaction on Image Processing 21, pp. 4695–4708, (2012)
[21] Mittal, A., Soundararajan, R., Bovik, A. C., "Making a Completely Blind Image Quality Analyzer." IEEE Signal Processing Letters, 22(3), pp. 209–212, (2013)
[22] Zhang, L., Zhang, L., Mou, X., Zhang, D., "FSIM: A feature similarity index for image quality assessment," IEEE Transactions on Image Processing, 20(8), pp. 2378–2386, (2011)
[23] Shahid, M., Rossholm, A., et al, "No-reference image and video quality assessment: a classification and review of recent approaches," EURASIP Journal on Image and Video Processing, 2014:40, (2014)
[24] Wu, H.R, Rao, K.R., "Digital Video Image Quality and Perceptual Coding (Signal Processing and Communications)," CRC, Boca Raton, (2005)
[25] Hemami SS, Reibman AR: No-reference image and video quality estimation: applications and human-motivated design. Signal Process. Image Commun., 25(7), pp. 469-481, (2010)
[26] Lin, W., Kuo, J.C.C., "Perceptual visual quality metrics: a survey." J. Vis. Commun. Image Representation, 22(4), pp. 297-312, (2011)
[27] Chandler, D.M., "Seven challenges in image quality assessment: past, present, and future research," ISRN Signal Processing, 53, (2013)
[28] Pedersen, M., Hardeberg, J.Y., "Full-reference Image Quality Metrics: Classification and Evaluation," Foundations and Trends® in Computer Graphics and Vision, 7(1), pp. 1-80, (2012)
[29] Avcibas, I., Sankur, B., Sayood, K., "Statistical evaluation of image quality measures," Journal of Electronic Imaging, 11, pp. 206–223, (2002)
[30] Le Callet, P., Barba, D., "A robust quality metric for color image quality assessment," in International Conference on Image Processing (ICIP), 1, pp. 437–440, (2003)
[31] Wang, Z., Bovik, A.C., "Modern Image Quality Assessment," Morgan & Claypool Publishers, (2006)
[32] Chandler, D.M., Hemami, S.S., "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," IEEE Transactions on Image Processing,16(9), pp. 2284–2298, (2007)
[33] Thung, K.H., Raveendran, P., "A survey of image quality measures," in International Conference for Technical Postgraduates (TECHPOS), pp. 1–4, (2009)
[34] Seshadrinathan, K., Bovik, A. C., "The encyclopedia of multimedia," in Chapter Image and Video Quality Assessment, Springer-Verlag, pp. 8–17, (2009)
[35] Kuryati, K., Shankar, K., et al.," Full Reference Image Quality Metrics and their Performance," IEEE 7th International Colloquium on Signal Processing and its Applications, pp. 34-38, (2011)